arXiv:quant-ph/0003040v1 13 Mar 2000

Unified approach to quantum capacities: towards quantum noisy coding theorem

Michał Horodecki[1], Paweł Horodecki[2] and Ryszard Horodecki [1]
[1] *Institute of Theoretical Physics and Astrophysics, University of Gdańsk, 80–952 Gdańsk, Poland,*
[2] *Faculty of Applied Physics and Mathematics, Technical University of Gdańsk, 80–952 Gdańsk, Poland*

Basing on unified approach to *all* kinds of quantum capacities we show that the rate of quantum information transmission is bounded by the maximal attainable rate of coherent information. Moreover, we show that, if for any bipartite state the one-way distillable entanglement is no less than coherent information, then one obtains Shannon-like formulas for all the capacities. The inequality also implies that the decrease of distillable entanglement due to mixing process does not exceed of corresponding loss information about a system.

Pacs Numbers: 03.65.-w

The challenge for the present quantum information theory is to determine the quantum capacity of noisy channels [1–7]. The problem is difficult mainly for two reasons. First, according to the present knowledge, unlike in the classical case, there are at least five different types of quantum capacities [1]. This is because quantum information channel can be supplemented by one- or two-way classical channel [8]. Moreover, there are teleportation channels [9,10], for which a bipartite state is a resource. The second reason is that quantum capacity exhibits a kind of nonadditivity [11] that makes them extremely hard to deal with.

As one knows, the key success of classical information theory is the famous Shannon noisy coding theorem, giving the formula for capacity of noisy channel [12]

$$C = \sup_X I(X;Y) \tag{1}$$

where the supremum is taken over all sources X; $I(X;Y) = H(X) + H(Y) - H(X;Y)$ is the Shannon mutual information (with H being the Shannon entropy); Y is the random variable resulting from action of the noise to X. In quantum information theory the candidate for the counterpart of mutual information has been found [2,3]. It is the so-called *coherent information* (CI). To define it, consider two coherent informations of the bipartite state ϱ with reductions ϱ_A and ϱ_B:

$$I^X(\varrho) = S(\varrho_X) - S(\varrho), \quad X = A, B, \tag{2}$$

for $S(\varrho^X) - S(\varrho) \geq 0$ and $I^X = 0$ otherwise. Here $S(\varrho) = -\mathrm{Tr}\varrho \log_2 \varrho$ is the von Neumann entropy. Then one defines the coherent information for a channel Λ and a source state σ as

$$I(\sigma, \Lambda) = I^B((I \otimes \Lambda)(|\psi\rangle\langle\psi|)), \tag{3}$$

where ψ is a pure state with reduction σ (the quantity does not depend on the choice of ψ). Now, the following connection between coherent information and a quantum capacity is known [7]

$$Q_\emptyset \leq \lim_{n\to\infty} \frac{1}{n} \sup_{\varrho_n} I(\varrho_n, \Lambda^{\otimes n}) \equiv I^B_\emptyset(\Lambda) \tag{4}$$

where $Q_\emptyset$ is the maximal number of qubits that can be reliably sent down the channel without any supplementary classical channel. Note that if, instead of inequality, there were *equality*, then we would have analogue of Shannon formula.

Unfortunately, despite a huge effort devoted to the problem [4,6,7], the equality has not been proven so far. Moreover, remarkably, the similar inequality is not known for other capacities than $Q_\emptyset$, i.e. the ones attainable at the support of backword ($Q_\leftarrow$) or two-way ($Q_\leftrightarrow$) classical communication [13]. There are also capacities of quantum teleportation channels, where the resource is bipartite state rather than channel. In the latter case, transmission requires prior manipulations (called distillation [14,1]) over the shared pairs, transforming them into pairs in pure maximally entangled states. Then the quantum information can be transmitted by using teleportation [9]. The manipulations include any local actions, and one- or two-way classical communication. Correspondingly, we have two kinds of one-way distillable entanglement of a state ϱ, $D_\rightarrow$ or $D_\leftarrow$ (since ϱ need not be symmetric, one distinguishes directions of classical communication) and two-way distillable entanglement $D_\leftrightarrow$. In fact, $Q_\leftrightarrow$ also necessarily involves the distillation and teleportation process.

The latter processes are so exotic from the point of view of classical information theory, that no analogue of theory of error correcting codes has been worked out for them so far. In contrast, in the case of the capacity $Q_\emptyset$, there exists a huge theory of quantum codes, being a generalization of classical error correcting codes theory [15].

In the above context the basic question arise: *Is there possible a consistent approach for all of the capacities*? In particular: *Is there a single counterpart of the Shannon mutual information*? In this paper, we provide a *unified* framework for *all* capacities. We show that one and the same coherent information, although in different contexts, is a basic quantity in each case. More specifically, we show that the inequality (4) is, in a sense, *universal*. Given any type of supplementary resources, the maximal rate of quantum communication (quantum capacity) is bounded by the maximal rate of coherent information attainable via these resources (CI capacity).

Now, there remains the fundamental question: Are the quantum capacities equal to corresponding CI capacities? We will show that the following hypothetical inequality (call it *hashing inequality* [16])

$$D_{\rightarrow}(\varrho) \geq I^B(\varrho), \tag{5}$$

if satisfied for all bipartite states ϱ, implies that the capacities are equal to one another in all cases (see Theorem 2). In other words, the hashing inequality implies the Shannon-like formulas for quantum capacities, providing the quantum noisy coding theorem. Consequently, we argue that to prove (or disprove) this inequality is one of fundamental tasks of the present information theory. In particular, if the inequality holds, then to evaluate $Q_{\emptyset}$ one would need to consider only the maximization problem of the right-hand-side of the inequality (4). Finally, we show that the power of the above inequality is even more surprising. Namely, it also implies the relation between loss of classical information, and loss of distillable entanglement [17].

Surprisingly, the reasoning leading to our results on capacities is extremely simple. Namely, the capacity of a channel or bipartite state at given supplementary resource is the optimal rate of reliable transmission of qubits. However, this is equivalent to optimal rate of reliable sharing two-qubit pairs in maximally entangled (in short, singlet) states [1]. Thus, we can imagine that Alice and Bob, started with large number n of pairs in initial state $\varrho^{\otimes n}$ (or disposed n uses of quantum channel Λ), aim to share the maximal attainable number k of singlet pairs. The capacity is just the optimal rate k/n. Then, what is the coherent information I^X_{out} of the output of such protocol attaining capacity? Since the coherent information is additive, and for singlet state $I^X = 1$, then I^X_{out} equals to the number k of final singlet pairs. Thus the obtained *rate* of the coherent information I^X_{out}/n is equal to capacity. But the *maximal* attainable rate of coherent information (i.e. CI capacity) is no less than the one acheievable in some particular protocol, so that it is no less than the capacity.

Assume now, that for any bipartite state, its capacity (i.e. distillation rate) is no less than its coherent information. Then we consider the following protocol. Alice and Bob start from k groups of n pairs (or divide kn uses of channel into k groups), with k, n being large, and for any group obtain the final state ϱ of maximal attainable coherent information. Then they distill the latter state, and, by assumption, obtain the final number of singlet pair no less than the coherent information of ϱ. Since the latter was maximal attainable one, we conclude that the capacity of the input state ϱ or channel Λ is no less than CI capacity. Therefore, due to previous paragraph, the quantities must be equal.

The presented argumentation is very intuitive. It is similar to the approach of Ref. [19], that has already appeared to be fruitful in a different context [20]. The rigourous version of the above heuristic approach is more or less immediate. Indeed, the main simplification we made was the assumption that *exact* singlets are produced. In fact, they are always impure (the impurity vanishes in the asymptotic or "thermodynamic" limit of infinite number of input pairs). However such simplification does not lead to wrong conclusions, if only the involved functions exhibit suitable continuity. In the rigorous proofs below we will use continuity of coherent information.

Let us now pass to the rigorous part of the paper. As mentioned, we will be concerned with four supplementary resources $C \in \{\rightarrow, \leftarrow, \leftrightarrow, \emptyset\}$. (The last one symbolizes no supplementary resource). If Alice and Bob dispose one use of a channel Λ (directed, by convention, from Alice to Bob) and the supplemetary resources symbolized by C, then they can share a bipartite state ϱ. An operation that produced in this way the state ϱ from Λ will be denoted by $\mathcal{E}_C$ so that

$$\varrho = \mathcal{E}_C(\Lambda). \tag{6}$$

If Alice and Bob share initially a bipartite state ϱ_{in}, then we will use notation $\mathcal{D}_C$

$$\varrho_{out} = \mathcal{D}_C(\varrho_{in}) \tag{7}$$

(The letters used in our notation follows from the common associations: usual channel capacity – error correction, teleportation channels – distillation). Now, the CI capacitites are defined by

$$I^X_C(\Lambda) = \lim_{n\to\infty} \frac{1}{n} \sup_{\mathcal{E}_C} I^X(\mathcal{E}_C(\Lambda^{\otimes n})) \tag{8}$$

for channels, and

$$I^X_C(\varrho) = \lim_{n\to\infty} \frac{1}{n} \sup_{\mathcal{D}_C} I^X(\mathcal{D}_C(\varrho^{\otimes n})) \tag{9}$$

for bipartite states [21]. Throughout the paper, the symbol X stands for A or B.

We also define quantum capacities as follows [22]. Define maximally entangled state on the space $\mathcal{H} \otimes \mathcal{H}$ by

$$P_+(\mathcal{H}) = |\psi_+(\mathcal{H})\rangle\langle\psi_+(\mathcal{H})|, \quad \psi_+(\mathcal{H}) = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |ii\rangle, \tag{10}$$

where $|i\rangle$ are basis vectors in $\mathcal{H}$, while $d = \dim \mathcal{H}$. Given a state ϱ, consider sequence of operations $\{\mathcal{D}^n_C\}$ (called protocol) transforming the input state $\varrho^{\otimes n}$ into the state σ_n acting on the Hilbert space $\mathcal{H}_n \otimes \mathcal{H}_n$ and with $\dim \mathcal{H}_n = d_n$, satisfying

$$F_n \equiv \langle\psi_+(\mathcal{H}_n)|\sigma_n|\psi_+(\mathcal{H}_n)\rangle \to 1. \tag{11}$$

The asymptotic ratio attainable via given protocol is then given by

$$D_{\{\mathcal{D}_C^n\}}(\varrho) = \lim_{n\to\infty} \frac{\log_2 \dim \mathcal{H}_n}{n} \tag{12}$$

Then the capacity $D_C(\varrho)$ (call it C-distillable entanglement) is defined by maximum over all possible protocols

$$D_C(\varrho) = \sup D_{\{\mathcal{D}_C^n\}}(\varrho). \tag{13}$$

The usual channel capacities can be defined in the same way. We only need make the following substitutions: $D \to Q$, $\varrho \to \Lambda$ and $\mathcal{D} \to \mathcal{E}$. The protocols $\{\mathcal{E}_C^n\}$ and $\{\mathcal{D}_C^n\}$ that achieve the considered suprema will be called optimal error correction and optimal distillation protocol, respectively. The quantity $D_\varnothing$ is a bit pathological, but certainly interesting quantity. We will not be concerned with it here. However, it is likely, that $D_\varnothing$ is the amount of pure entanglement that can be drawn from the state reversibly.

We will need a lemma, stating that coherent information I^X is continuous on isotropic state. The latter is defined on $\mathcal{H}\otimes\mathcal{H}$ (cf. [10,23,24])

$$\varrho(F,d) = pP_+(\mathcal{H}) + (1-p)\frac{1}{d^2}I, \quad 0 \le p \le 1, \tag{14}$$

with $\mathrm{Tr}\big[\varrho(F,d)P_+(\mathcal{H})\big] = F$, $d = \dim\mathcal{H}$.

Lemma. For a sequence of isotropic states $\varrho(F_n, d_n)$, such that $F_n \to 1$ and $d_n \to \infty$ we have

$$\lim_{n\to\infty} \frac{1}{\log_2 d_n} I^X(\varrho(F_n,d_n)) \to 1. \tag{15}$$

Proof. This can be checked by direct calculation.

We note an important property of the isotropic state [24]. Namely, any state σ acting on $\mathcal{H}\otimes\mathcal{H}$, if subjected to $U\otimes U^*$ twirling (cf. [14]), i.e. random unitary transformations of the form $U\otimes U^*$, becomes isotropic state $\varrho(d,F)$ with $F = \mathrm{Tr}\big[\sigma P_+(\mathcal{H})\big]$, $d = \dim\mathcal{H}$.

Now we can state the theorems being the main results of this paper.

Theorem 1. Quantum capacities are bounded from above by CI capacities:

$$Q_C(\Lambda) \le I_C^X(\Lambda), \tag{16}$$
$$D_C(\varrho) \le I_C^X(\varrho) \tag{17}$$

for any Λ, ϱ and $C \in \{\to,\leftarrow,\leftrightarrow,\varnothing\}$.

Theorem 2. If the hashing inequality

$$D_\to(\varrho) \ge I^B(\varrho) \tag{18}$$

holds for any bipartite state ϱ then the qantum capacities are equal to corresponding CI capacities

$$Q_\leftrightarrow(\Lambda) = I_\leftrightarrow^X(\Lambda), \quad Q_{\underset{(\leftarrow)}{\to}}(\Lambda) = I_{\underset{(\leftarrow)}{\to}}^{B(A)}(\Lambda), \tag{19}$$
$$Q_\varnothing(\Lambda) = I_\varnothing^B(\Lambda), \tag{20}$$
$$D_\leftrightarrow(\varrho) = I_\leftrightarrow^X(\varrho), \quad D_{\underset{(\leftarrow)}{\to}}(\Lambda) = I_{\underset{(\leftarrow)}{\to}}^{B(A)}(\varrho). \tag{21}$$

Remarks. (i) If we assumed "dual" hashing inequality $D_\leftarrow \ge I^B$, we would get the same results modulo change $A \leftrightarrow B$. Our choice of $D_\to \ge I^B$ is motivated by investigations of Refs. [2–4]. (ii) It follows that the hashing inequality implies $I_\leftrightarrow^A = I_\leftrightarrow^B$. (iii) Our results apply to other kind of supplementary resurces such as e.g. public bound entanglement.

Proof of Theorem 1. We will prove the "Q" part of the theorem. The proof for "D" part is similar. Let $\{\mathcal{E}_C^n\}$ be the optimal error correction protocol for Λ. Then we have the following estimates

$$I_C^X(\Lambda) \ge \frac{1}{n} I^X(\mathcal{E}_C^n(\Lambda^{\otimes n})) \ge \frac{1}{n} I^X(\varrho(d_n,F_n)) \to Q_C(\Lambda), \tag{22}$$

where $\varrho(d_n,F_n)$ is the twirled state $\sigma_n = \mathcal{E}_C^n(\Lambda^{\otimes n})$. The first inequality comes from the very definition of I_C^X. The second one follows from convexity of I^X [4], and its invariance under product unitary tranformations (as the twirled state is a mixture of product unitary tranformations of the initial one). Finally, since in optimal error correction protocol we have $F_n \to 1$, and $\log d_n/n \to Q_C$, we obtain the right-hand-side limit by applying continuity of I_X (see lemma).

Proof of Theorem 2. We will also prove only the "Q" part. For $C \in \{\to,\leftarrow,\leftrightarrow\}$, consider the following particular error correcting protocol for the channel $\Gamma = \Lambda^{\otimes n}$. One applies to Γ the operation $\mathcal{E}_C$ that produces the state $\sigma = \mathcal{E}_C(\Lambda^{\otimes n})$ of maximal attainable coherent information. Subsequently, one performs optimal distillation protocol for the state σ. Then we find

$$\begin{aligned} Q_C(\Lambda) = \frac{1}{n} Q_C(\Lambda^{\otimes n}) \ge \frac{1}{n} D_C(\mathcal{E}_C(\Lambda^{\otimes n})) \ge \\ \frac{1}{n} I^X(\mathcal{E}_C(\Lambda^{\otimes n})) \to I_C^X(\Lambda) \end{aligned} \tag{23}$$

where $C \in \{\to,\leftarrow,\leftrightarrow\}$; $X = A, B$ for $C = \leftrightarrow$, and $X = A\,(B)$ for $C = \leftarrow (\to)$. The equality follows from the very definition of Q_C. The first inequality comes from the fact, that Q_C is supremum over *all* error correction protocols, so it is no less from the rate obtained in the protocol above. The second inequality follows from the hashing inequality, and from the obvious inequality $D_\leftrightarrow \ge D_{\underset{(\leftarrow)}{\to}}$. Finally, the limit is due to the definition of $I_C^X(\Lambda)$. The above estimate together with Theorem 1 gives all the desired equalities apart from the one involving $C = \varnothing$. That the latter one is also implied by the hashing inequality, it follows immediately from the facts: (a) trivially $I_\varnothing^B \le I_\to^B$; (b) $Q_\to = Q_\varnothing$ [7]; (c) as just proved, the hashing inequality implies $Q_\to = I_\to^B$.

Let us now prove yet another important implication of the hashing inequality. Namely, consider the process of discarding information

$$\{p_i, \varrho_i\} \to \varrho = \sum_i p_i \varrho_i. \tag{24}$$

In Ref. [17] it was shown that, for a class of ensembles $\{p_i, \varrho_i\}$, the amount of information lost in the process is no less than the loss of distillable entanglement $D_\leftrightarrow$, and it was conjectured to hold in general. The loss of information is quantified by average increase of entropy, so that the problem is whether the folowing inequality holds

$$\sum_i p_i D_C(\varrho_i) - D_C(\varrho) \le S(\varrho) - \sum_i p_i S(\varrho_i). \tag{25}$$

Note, that for pure state ψ, $D_C(\psi) = S(\varrho^X)$, where ϱ^X is either of the reductions of ψ [25]. Therefore, for pure states ϱ_i, the inequality reads

$$D_C(\varrho) \ge \sum_i p_i S(\varrho_i^X) - S(\varrho) \tag{26}$$

Applying convexity of entropy we see that the hashing inequality implies the above one. It is interesting, that it does not seem to imply the inequality for impure ϱ_i's.

Let us list that recent results concerning entanglement distillation, implying that it is reasonable to conjecture that the hashing inequality (5) holds. (i) In all cases where one has sufficiently tight lower bounds for $D_\to$, the inequality is known to be satisfied. For pure states, and other ones with entanglement of formation equal to entanglement of distillation [26] we have $D_\to = I^B$. For mixtures of two-qubit Bell states we have $D_\to \ge I^B$ by hashing protocol [1]. In particular, for some of them there is equality [27], while for other ones one has $D_\to > I^B$ [11]. (ii) If the hashing inequality is true, then any upper bound for $D_\leftrightarrow$ should be no less than I^X. This was shown for entanglement of formation [26] and, quite recently, for relative entropy of entanglement [28]. We do not know yet, if the inequality holds for the new bound for D derived in [20]. (iii) If a state is bound entangled [29], then we should have $I^X = 0$. It is indeed the case. According to Ref. [24], the bound entangled states must satisfy the so called *reduction criterion of separability.* This implies [30] that the entropic inequality $S(\varrho) \ge S(\varrho_X)$ is also satisfied, hence $I^X = 0$. Thus we see that there is a strong evidence that the inequality is true. We believe that the present results will stimulate to prove (or disprove) it.

The work is supported by Polish Committee for Scientific Research, contract No. 2 P03B 103 16, and by the IST project EQUIP, contract No. IST-1999-11053.

---

[1] C. H. Bennett, D. P. Di Vincenzo, J. Smolin and W. K. Wootters, Phys. Rev. A **54**, 3814 (1997).

[2] B. Schumacher, Phys. Rev. A **54**, 2614 (1996); B. Schumacher and M. A. Nielsen, Phys. Rev. A **54**, 2629 (1996).

[3] S. Lloyd, Phys. Rev. A **56**, 1613 (1997).

[4] H. Barnum, M. Nielsen and B. Schumacher Phys. Rev. A **57**, 4153 (1998).

[5] C. H. Bennett, D. Di Vincenzo and J. Smolin, Phys. Rev. Lett. **78**, 3217 (1997).

[6] H. Barnum, J. Smolin, and B. Terhal Phys. Rev. A **58**, 3496 (1998).

[7] H. Barnum, E. Knill and M. Nielsen, quant-ph/9809010.

[8] There are also *classical* capacities of quantum channels, see B. Schumacher and M. Westmoreland Phys. Rev. A **56**, 131 (1997); A. S. Holevo, IEEE Trans. IT **44**, 269 (1998); Bennett *et al.* Phys. Rev. Lett. **83**, 3081 (1999); Here we are concerned with solely *quantum* capacities.

[9] C. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres and W. K. Wootters, Phys. Rev. Lett. **70** (1993) 1895.

[10] S. Popescu, Phys. Rev. Lett. **72**, 797 (1994).

[11] P. Shor and J. Smolin, quant-ph/9706061; D. DiVincenzo, P. Shor and J. Smolin Phys. Rev. A **57**, 830 (1998).

[12] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley and Sons, New York, 1991.

[13] In Ref. [7] it was shown that forward classical communication does not increase capacity, i.e. $Q_\emptyset = Q_\to$. The interesting question is if $Q_\leftarrow = Q_\emptyset$.

[14] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. Smolin and W. K. Wootters, Phys. Rev. Lett. **76**, 722 (1996).

[15] P. Shor, Phys. Rev. A, **52**, 2439 (1995); A. Steane, Phys. Rev. Lett. **77**, 793 (1996); Th. Beth and M. Grassl, Fortschr. Phys. **46**, 459 (1998) and references therein.

[16] For a class of states, the inequality was shown to be true by use of the so-called *hashing* distillation protocol [1].

[17] J. Eisert, T. Felbinger, P. Papadopoulos, M. B. Plenio and M. Wilkens, Phys. Rev. Lett. **84**, 1611 (2000).

[18] We will assume, without explicit stating it each time, that a fixed supplementary resource is used.

[19] M. Horodecki, P. Horodecki and R. Horodecki, Phys. Rev. Lett. **84**, 2014 (2000).

[20] M. Horodecki, P. Horodecki and R. Horodecki, Report No. quant-ph/9912076.

[21] By convexity of I^X [4] this definition is consistent with notation used in formula (4).

[22] The used definition of "D" capacities is due to [27] (cf. [1]). In Ref. [E. Rains, Phys. Rev. A **60**, 173 (1999)] the definition was shown to coincide with other used definitions. The definition of "Q" capacities is taken from [5]. In Ref. [7] it was shown to be equivalent with other one introduced in [4].

[23] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).

[24] M. Horodecki and P. Horodecki, Phys. Rev. A, **59** 4206 (1999).

[25] C. H. Bennett, H. J. Bernstein, S. Popescu and B. Schumacher, Phys. Rev. A **53**, 2046 (1996).

[26] P. Horodecki, M. Horodecki and R. Horodecki, Acta Phys. Slovaca **48**, 141 (1998).

[27] E. M. Rains, Phys. Rev. A **60**, 179 (1999).

[28] M. B. Plenio, S. Virmani, P. Papadopoulos, Report No. quant-ph/0002075.

[29] M. Horodecki, P. Horodecki and R. Horodecki, Phys. Rev. Lett. **80**, 5239 (1998).
[30] N. Cerf, C. Adami and R. M. Gingrich, Phys. Rev. **60**, 898 (1999).